DO "HYPER-TRAVELLERS" EXIST? – INITIAL RESULTS OF AN INTERNATIONAL SURVEY ON PUBLIC TRANSPORT USER BEHAVIOUR

Achille Fonzone, Imperial College London, a.fonzone@imperial.ac.uk

Jan-Dirk Schmöcker, Kyoto University, schmoecker@trans.kuciv.kyoto-u.ac.jp

Michael G. H. Bell, Imperial College London, m.g.h.bell@imperial.ac.uk

Guido Gentile, Università di Roma "La sapienza", guido.gentile@uniroma1.it

Fumitaka Kurauchi, Gifu University, kurauchi@gifu-u.ac.jp

Klaus Nökel, PTV, Klaus.Noekel@ptv.de

Nigel H. M. Wilson, Massachusetts Institute of Technology, nhmw@mit.edu

ABSTRACT

Transport modelling – both in traffic and transit fields – has been traditionally based on the assumption of the utility maximization principle. In public transport networks with high frequency services this has led to "hyperpath" route choice models in which passengers reduce their travel time by identifying for each node a set of attractive lines, each of which might be the fastest from the node, depending on its and other lines arrival time. Several assignment models have been developed using hyperpaths but the question of "whether" and "to what extent" transit users are able to reason and act in terms of hyperpaths has not yet been addressed. Doubts generated by the research on route choice call for experimental studies of the issue. The paper reports the initial results of a web-based survey of public transport frequent users in six countries (China, Germany, Japan, Italy, UK, and USA). Analyses of the replies concerning travellers' actual behaviour are presented, which provide insights on the flexibility of route choice in relation to trip characteristics and information on services. Further behavioural research is identified which can lead to improved transit route choice and assignment models.

Keywords: Public transport; Traveller behaviour; Hyperpath models; Travel information; Web-based survey

INTRODUCTION

Most of route choice models assume travellers to follow a utility maximizing behaviour with cost functions coinciding or at least including expected trip times. In networks characterized by randomness the expected trip time can be reduced by an adaptive behaviour, i.e. by subordinating the route choice to the events which actually occur during the journey. The result has been proved by Hall (1986) for networks with stochastic link travel times, and by Spiess and Florian (1989) for transit networks in which stochasticity is related to the order of arrival of buses at a stop. In the latter passengers are supposed to follow strategies of the kind "Take the line which arrives first among those in your attractive choice set". Attractive sets are defined by minimizing an expected trip time, sum of the waiting times at the nodes plus the travel times along the links. Such behaviour can be represented by hyperpaths (Nguyen and Pallottino, 1988), i.e. paths between OD pairs in which each node can have more than one successor or, equivalently, in which at each node more than one link can exist with a positive probability of being used.
Literature especially considering the "bus stop problem" has proposed corrections to the simple Spiess and Florian model, in case bus arrival times cannot be assumed to be exponentially distributed (see Gentile et al., 2005; Nokel and Wekeck, 2009). Further several extensions have been proposed that widen the hyperpath set by taking into account additional sources of uncertainty, such as capacity problems (e.g. De Cea and Fernandez, 1993), seat availability (Schmöcker et al., 2009) and bus-bunching (Shimamoto et al., 2010).
Adaptive strategies seem to provide a realistic approach for transit user behaviour modelling, above all when systems are characterized by frequent services and overlapping lines but also in some applications in the context of the schedule-based models (see for instance Hamdouch and Lawphongpanich, 2008). However the characteristics of choice sets depend on the assumptions concerning the rationality of users (Fonzone and Bell, 2010). In hyperpath models currently in use, travellers are supposed utility maximizers endowed with unbounded computing capacity and perfect – even if stochastic as to line arrivals at stops – information, i.e. they perceive correctly the topology of the entire network, the travel times on every link and the probability distribution functions of every line at every stop.
Geographic research has established that the spatial decision making is "boundedly rational" (Wolpert, 1964) and makes use of cognitive maps, that is mental representations of the external world as it is experienced by the subject (Golledge and Stimson, 1997). The spatial learning in route choice develops through a process of implementing and evaluating trial trips, whose output is travel habits. Once travel habits are acquired, consciousness plays a small role in decision making and choices tend to be rigid. Field and laboratory experiments – mostly concerning private transport – have found out a wide assortment of criteria used in route choice: minimization of cost, in terms of time, distance, and generalized cost; minimization of turns, obstacles, route segments and intermodal changes; avoidance of congestion and dangerous areas, longest leg first; restriction to well known corridors; maximizing aesthetics (see Golledge and Garling, 2001, who generally provide a good starting point on this topic from a behavioural research point of view). The shortest path is regularly adopted by 50% of people (Golledge, 1997), but criteria change with the trip purpose. Recently much effort has been put into understanding the role of dynamic

information in travel decision making (Lyons, 2006). Electronic fare collection and GPS systems allow tracking users and are leading to a larger use of revealed preferences in behaviour research (e.g., Papinski et al., 2009; Seaborn, 2009).
In light of the existing research on route choice the behavioural model underpinning the hyperpath approach in transit modelling seems rather naive. At the best of our knowledge, no research has been carried out so far to test it. This paper describes the first results from an exploratory study of the behaviour of public transport users in China, Germany, Italy, Japan, UK and USA, carried out through a web-based survey. The aim is to understand the flexibility of travellers in choosing their paths in relation to the information they use, and to evaluate the realism of the assumptions that underlie transit route choice models, in particularly those assuming travellers adopting adaptive strategies. The outcomes of the research are expected to lead to improvements in transit assignment models.
The remainder is organized as follows: A description of the survey is provided in the next paragraph. SAMPLE presents the main features of the respondents, discussing issues related to the representativeness of the collected sample. ACTUAL BEHAVIOUR describes the characteristics of the trips made by respondents with a focus on the flexibility of choices, and the information used for trip planning; the attitude to introduce changes in routine itineraries is analysed through logistic and categorical regressions. DISCUSSION OF RESULTS provides a reading of the results in relation to transit modelling. CONCLUSIONS bring forward further research which can lead to more realistic transit route choice models.

SURVEY

In order to reach a large number of people in geographically distant places, and to allow for sufficient time for respondents to answer the numerous and not always simple questions, a web-based survey was developed. Potential respondents have been contacted principally by email. The main, but not exclusive, distribution channels were mailing lists of engineering students and transport specialists. Responses were collected between November 2009 and January 2010.
The questionnaire[1] is made up of three sections and 36 questions, as described in Table I. "Personal information" concerned age, gender, working status as well as place where respondents live and study. In the section "Actual behaviour" respondents were asked to consider a trip by public transport they frequently make. Then firstly characteristics of these trips were asked for such as time duration, public transport means used or whether the trip requires interchanging. Respondents were further asked to answer questions on the information sources they use to plan the trip and potentially inform themselves about alternatives. To understand route choice flexibility in particular respondents were further asked to state whether they do consider alternative routes by varying for example their departure station or route choice from their departure or an interchange station. The third part of the questionnaire on "Hypothetical route choice scenarios" is not discussed in this paper. The answer to some questions was conditional on the answer to the previous ones, i.e. respondents might be advised to skip these questions. Hypothetical scenarios are

[1] Available at https://academictrial.qualtrics.com/SE?SID=SV_096hoL4rFUU9cDa&SVID=

depicted to respondents through simplified maps. The time spent by respondents on each survey section was recorded in the background. Besides the English version, translations in Chinese, German, Italian and Japanese were provided.
The survey was approached 1,022 times and was completed (i.e., the button "Submit" at the end of the questionnaire was pressed) in 579 cases. The 5% trimmed mean of the survey completion time is 20.5 min, with the 5^{th} percentile equal to 8.0 min and the 95^{th} to 121.0 min.

Table I – Structure of the questionnaire

Questionnaire structure						
		Questions				
Section	Subsection	Text entry	Multiple choice, single answer	Multiple choice, multiple answers	Matrix Table	Total
Personal information		3	2	0	1	6
Actual behaviour	Trip characteristics	4	4	1	1	10
	Available information	0	3	1	0	4
	Choice flexibility	0	4	4	0	8
Hypothetical route choice scenarios		0	7	0	1	8
Total		7	20	6	3	36

SAMPLE

Characteristics

The outcomes presented below are derived considering only the replies of the 579 respondents who completed the questionnaire. This is because 1) they are most likely to be motivated to participate in the survey and so to provide reliable answers; 2) taking into account also people who filled in only some of the questions would mean using non consistent samples for different statistics; 3) the regression analyses described later in this paper require a list-wise elimination of incomplete records and therefore incomplete questionnaires would not be included in any case.
38.0% of the respondents are women with a mean age of 29.6 years; and 90% are less than 42.0 years old. The male component of the sample has a mean age of 31.4 years and a 90^{th} percentile of the age distribution equal to 48.0 years (Figure 1a). The vast majority of respondents are either students or employees (Figure 1b).

(a) (b)

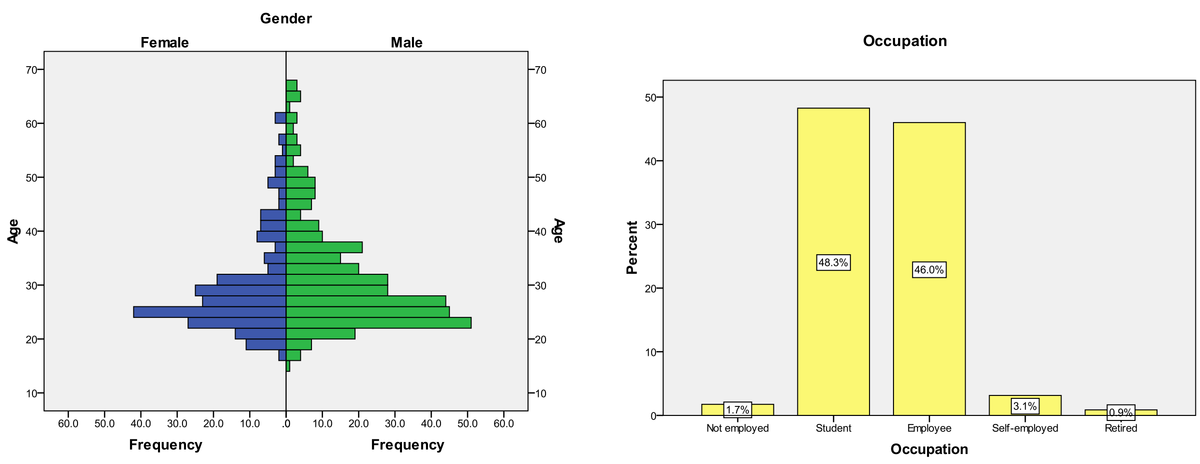


Figure 1 – (a) Age distribution per gender, (b) Occupational category

Participants come from 106 different work/study cities, which have been taken as reference to determine respondent's country and when geographical aspects are considered in the following; the 10 most represented cities are listed in Table II.

Table II – 10 most represented cities

City of job/study (first 10)				
City	Country	Overall percent	Overall cumulative percent	Percent within the country
London	UK	24.9	24.9	79.8
Roma	Italy	13.4	38.3	60.7
Tokyo	Japan	7.4	45.7	54.1
Karlsruhe	Germany	4.9	50.5	58.7
Taranto	Italy	4.5	55.1	25.0
Wuhan	China	4.3	59.4	46.2
Berkeley	USA	2.5	61.9	25.0
Graz	Austria	2.3	64.3	81.3
Kyoto	Japan	2.3	66.6	17.6
New York	USA	2.0	68.6	19.6

Replies arrive from 25 countries (namely Australia, Austria, Belgium, Canada, Chile, China, Czech, Denmark, France, Germany, Holland, Iran, Israel, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK, USA), with the six original target countries covering 90.5% of the sample. UK and Italy are more represented than other target countries (**Error! Reference source not found.**).

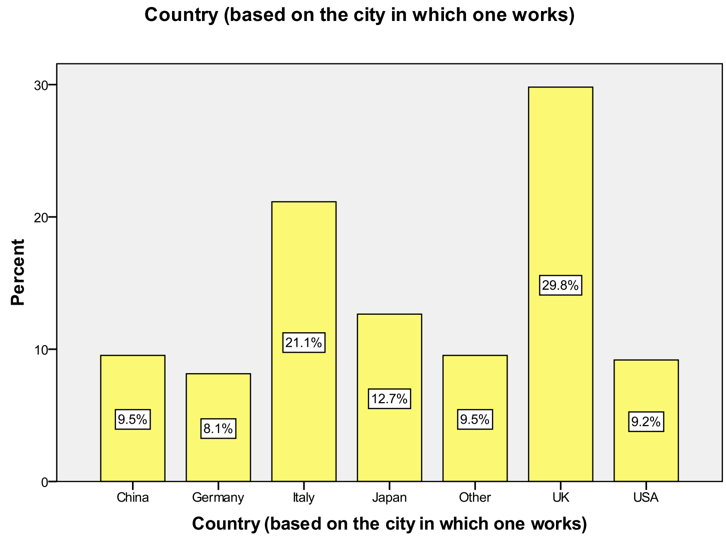


Figure 2 – Origin of the respondents

Respondents can be considered expert transit users: 70.0% travel by public transport 2-3 times a week or more.

Representativeness

The sample seems to be biased as to age, gender, and occupation of respondents. The choice of the web as platform for the survey could have brought forward a bias towards lower values in the age distribution. The gender split could have been influenced by the choice of the mailing lists addressed to distribute the survey: Most of them are in the engineering field, where, in some places, the male workforce is still predominant. Also the very low number of not employed, self employed and retired people is probably due to the way in which the survey was publicised. The lack of knowledge about the socio-demographics characteristics of public transport users in geographically and socially distant contexts such as those surveyed prevents from evaluating the representativeness of the sample, which in any case is extremely small compared to the whole population of the public transport users. However because of the exploratory nature of the survey it is deemed that even a sample not completely representative from the demographic point of view can grant useful results. This is somehow equivalent to assume that the behavioural characteristics we are interested in are not affected by demographics; consequently they have been not included in the models interpreting choice flexibility.
The high proportion is probably another bias of our sample, but it is intentional because our rationale is that, if even "experts" do not consider complex route choice strategies, occasional public transport users will even less. Travel behaviour and experience are strictly related, and both depend on the features of the transport system with which the user is familiar: e.g., it is reasonable to expect the travellers whose experience is limited to low frequency services or to systems with few overlapping lines to be less prone to consider multiple path alternatives in their decision making process, even though they are familiar with

public transport. This can be an issue when people are aggregated at a world level: Combining respondents with different experiences, without any kind of sample selection, might give rise to biased and difficult to interpret results. On the other side such a large geographic scale is helpful to capture behaviour invariants, if they exist, which is the aim of the present paper. A more detailed analysis of the effects of the origin of respondents will be addressed in future works.

ACTUAL BEHAVIOUR

Description

Participants were asked to describe their actual behaviour having in mind a specific trip they usually make by public transport. This is in order to limit the distortion of the collected information caused by the fact that users could have an incorrect perception of their own actions; indeed such a distortion can be amplified when information on average behaviour is asked.

Trip characteristics

2 out of 3 participants report commuting trips, the remaining third is equally distributed between business trips (both work-related and personal/family-related) and travelling for other activities (e.g., sport, leisure, visiting). Respondents were asked how important punctuality is for this trip. On a 5 point scale, the importance is rated 4.0 on average. 82.7% of trips take place in a weekday, prevailingly in the morning, with 69.9% starting in the morning peak (7-9 am). The minimum, average and maximum durations of the trip were recorded. Considering the 5% trimmed means, trips take on average 48.3 min, but their duration can vary between a mean minimum of 40.0 and a mean maximum of 66.4 min. 80% of respondents use public transport means that usually run as frequency-based services such as bus and underground, and 47.6% use services that usually are schedule-based (Table III). 94.0% of the trips involve 2 changes or less [referred to as result 1 in the following], one third are made up of a single (motorized) segment. Medium levels of congestion are reported in half of the cases, difficulty in getting on board in 14.8%.

Table III – Means used for the trip

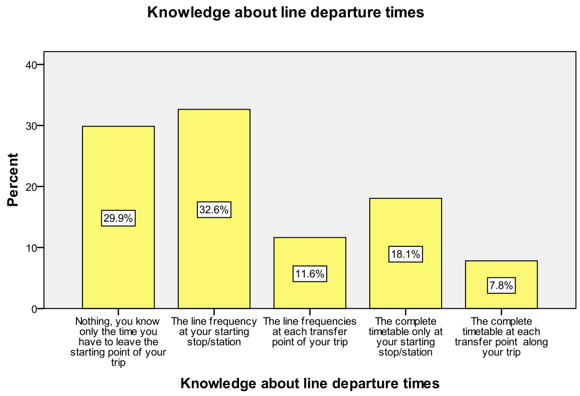
	Means combination			
	Train	Intercity bus	Urban bus / tram	Underground
Train	37.4%*	5.0%	13.8%	14.3%
Intercity bus	5.0%	15.2%	3.8%	5.4%
Urban bus / tram	13.8%	3.8%	51.2%	16.6%
Underground	14.3%	5.4%	16.6%	45.4%

* Percents on the main diagonal refer to the trips in which the concerned means is used, either alone or in combination with other listed means (e.g. a trip segment is covered by underground in 45.4% of the cases; 16.6% of trips – which are part of the 45.4% using the underground – combine the underground itself with urban bus / tram)

Information

Route choice models currently in use in the transit field assume that passengers know either the inter-arrival times (frequency-based models) or the timetables (schedule-based models) of each line of the network (or at least of the nodes which could be part of a path between the origin and destination of his trip). The survey shows that only 11.6% of the respondents know the frequency at each transfer point of the trip and just 7.8% the timetables (Figure 3a) [result 2]. In evaluating the results it has to be considered that they concern a trip with which respondents are probably well acquainted, given that they have decided to describe this trip in the survey. Half of the participants can anticipate the arrival time with a good precision once they have set the departure time (Figure 3b).

(a)



(b)

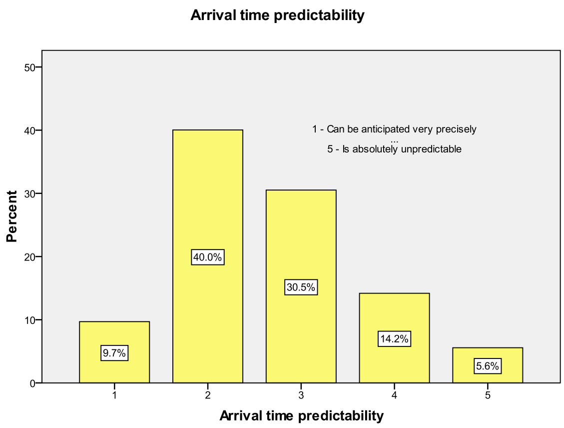


Figure 3 – Travellers' knowledge of the transit system about (a) service frequency and timetables, (b) arrival time

To plan their trip, more than 80% make use of information systems, either at home or during the journey (about 60% each, Figure 4a). When responses are analysed by the kind of systems and city the situation appears to be quite heterogeneous. This could be expected because in this case the behaviour depends on the supply of information systems, which varies between cities or even operators within cities (Figure 4b). Nevertheless some general conclusions can be drawn: at a global level and for all of the most represented cities of each country the most common sources of information are on-line journey planners at home, and timetables and displays at stops/stations (apart from Tokyo, where mobile phones are widely used) [result 3]. Journey planners used from home, which are generally very widespread, tend to be scarcely updated in their information and provide users with sets of single paths and not with adaptive strategies. Among systems which are usually updated in real time, displays are more used than mobile phones. The information coming from the former source does not allow completely rational traveller behaviour: In fact usually it concerns only a subset of the network (e.g. the lines serving a specific stop/station) or otherwise is purely qualitative (e.g. – for lucky people! – announcements like "A good service is currently operated on all London Underground lines"), hardly ever it is multimodal. Certainly the burden of adapting the routes when network conditions change is left to passengers. Where good services are offered, like in Japan or Berkeley, US, mobile phones are widely used.

(a)

Info at home * Info en-route Crosstabulation

% of Total

		Info en-route No	Info en-route Yes	Total
Info at home	No	17.1%	22.1%	39.2%
	Yes	22.5%	38.3%	60.8%
Total		39.6%	60.4%	100.0%

(b)

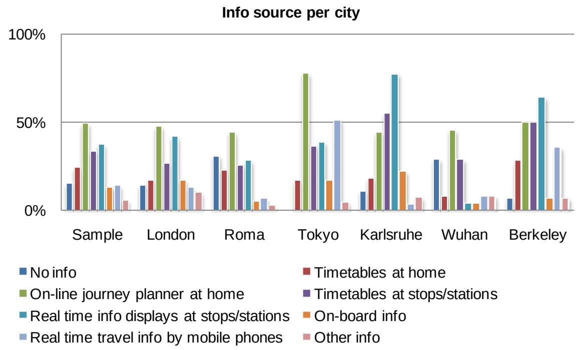
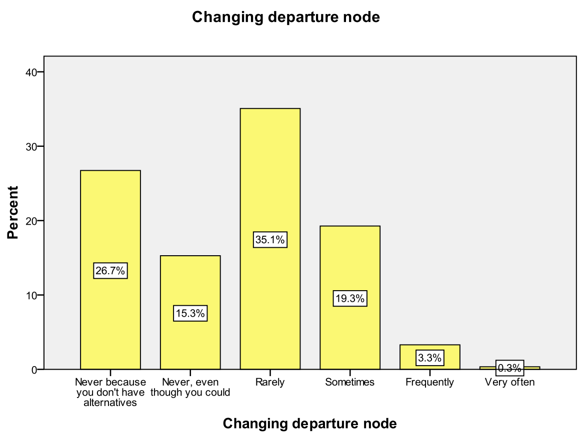


Figure 4 – Information systems – (a) At an aggregate level, (b) By kind and city

Flexibility of choices

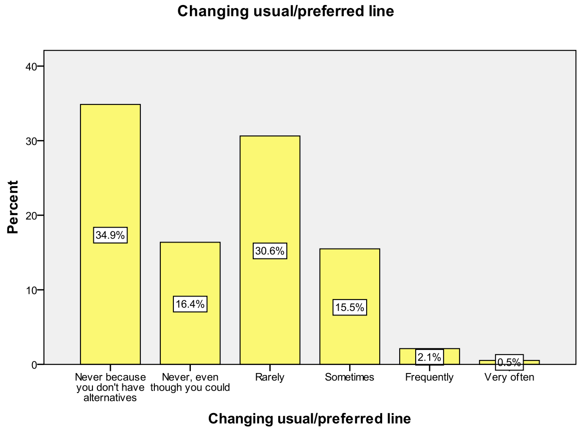
To evaluate whether public transport users are familiar with adaptive strategies, travellers' attitude towards choice flexibility has been investigated in relation to four dimensions. These are variation in choice of a) Departure stop/station, b) Usual/preferred lines, c) Transfer points and d) Alighting point once a passenger boarded a line.
In each case respondents have been asked to state how frequent their changes are, specifying in case they never change if this is due to a lack of alternatives or to a personal decision. Participants have also detailed the causes of their change in a multiple choice question. Results are reported in the four figures below, causes of changes are classified in related to level of service (Los), personal choice (Pc), external reason (Ex) or possibility to save time (Ts).



When you are about to start your trip (before going to the first stop/station), how often does it happen that you decide to leave from a stop/station different from the usual one?

Cause of change	
Service disruptions [*Los**]	31.0%
Additional purposes for the trip [*Pc*]	25.4%
Different departure time [*Pc*]	22.3%
Other	16.8%
The weather [*Ex*]	16.6%
Don't like always the same route [*Pc*]	12.1%

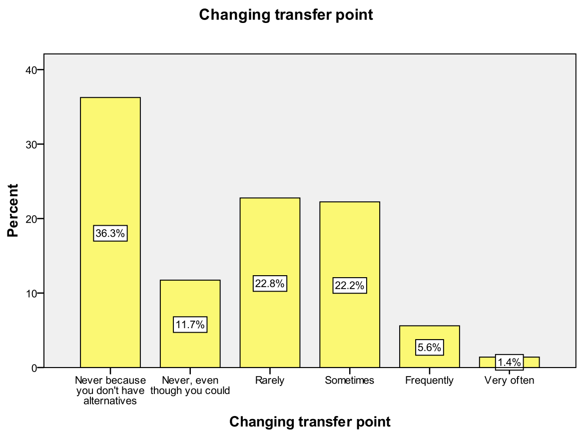
Figure 5 – Flexibility in changing departure stop/station



When you are at a stop/station, how often does it happen that you decide not boarding your usual/preferred line even though you're not forced to do so?

Cause of change	
Vehicle too packed [*Los*]	37.8%
Earlier arrival at destination [*Ts*]	34.6%
Cannot find a seat [*Los*]	11.4%
Other	10.0%
Don't like always the same route [*Pc*]	7.8%

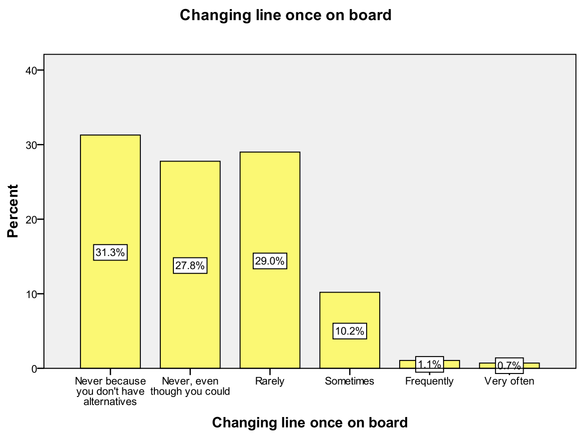
Figure 6 – Flexibility in changing usual/preferred lines



How often does it happen that, once you have alighted at an intermediate stop/station, you change to a transfer point different from the one you had in mind before alighting?

Cause of change	
Different useful line available [*Ts*]	31.3%
Transfer point too crowded [*Los*]	24.7%
Possibility of not getting on the next vehicle [*Los*]	21.4%
Other	15.4%
Possibility of not finding a seat on the next vehicle [*Los*]	12.4%
The weather [*Ex*]	11.5%

Figure 7 – Flexibility in changing transfer points



When you are on board, how often does it happen that you change line/vehicle even though you could reach your destination staying on the same vehicle?

Cause of change	
Earlier arrival at destination [*Ts*]	37.1%
Too crowded vehicle [*Los*]	21.5%
Bad operation [*Los*]	19.9%
Other	4.9%
Don't like always the same route [*Pc*]	4.3%

Figure 8 – Flexibility in changing lines once on board

All route choice aspects considered, 80.6% of the reported "typical" trips entail variations in their itineraries [result 4]. As it could be expected, in most cases (90.2%) participants have a choice available in at least one of the four dimensions, and they use it: The percentage of those who can modify their itineraries in at least one route choice aspect but never make it is just 11.3%; those who change "Sometimes" or more frequently are 63.9% [result 5]. Figure 9a summarizes the previous figures, highlighting that the most frequent type of change concerns the departure stop/station (almost 60% of cases considering at least "rarely"), the usual/preferred lines and the transfer points are changed more or less in half replies[2] [result 6]. Less frequent is the case in which a traveller gets off a service to take another one, even though he does not have to do so. When one focuses on the cases in which the users have choices, it appears that the propensity of users to change depends on the choice dimension (Figure 9b). Non parametric tests for more than two dependent samples confirm that the four samples in Figure 9b do not come from the same population: e.g. asymptotic significance = .000 in the Friedman test (Conover, 1999). Travellers seem to be more prone to change stops, stations and platforms than lines. On a 5-points scale of frequency, changing transfer points has been rated 2.41, changing departure point 2.16, changing usual/preferred line 2.08, and changing line once on board 1.81 [result 7]. For all dimensions, the most common response is "Rarely".

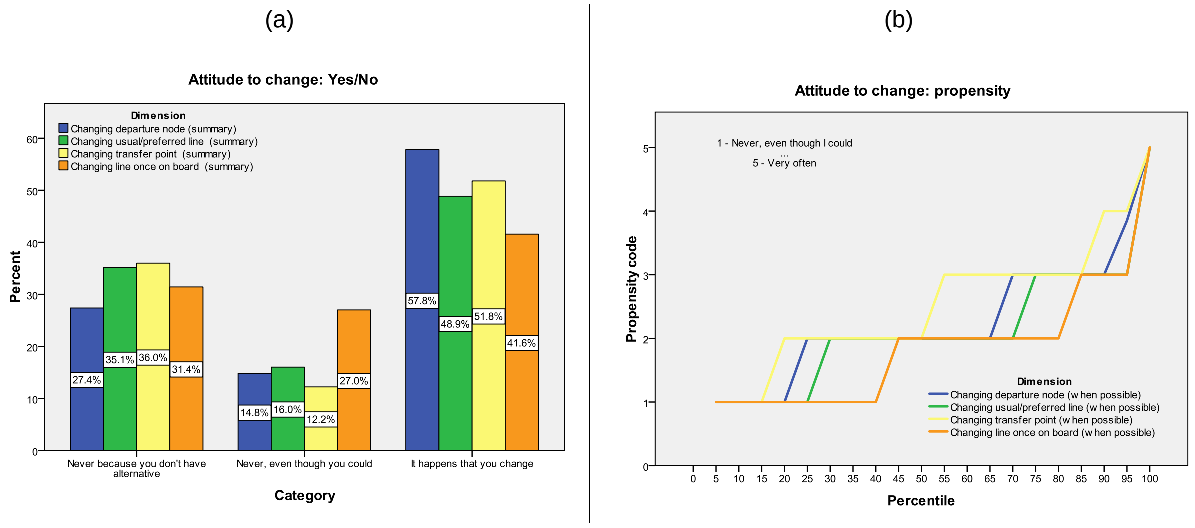


Figure 9 – Attitude to change – (a) Existence, (b) Propensity when possible

Figure 10a reports the occurrence of the different types of cause of change, assuming the classification of replies shown from Figure 5 to Figure 8. The possible causes of changes presented to users vary from question to question so relative importance cannot be easily inferred. A higher frequency seems to be associated with changes related to time saving (Figure 10b)

2 Numbers in Figure 9 are slightly different from those previously reported because of the list-wise selection of cases.

(a) (b)

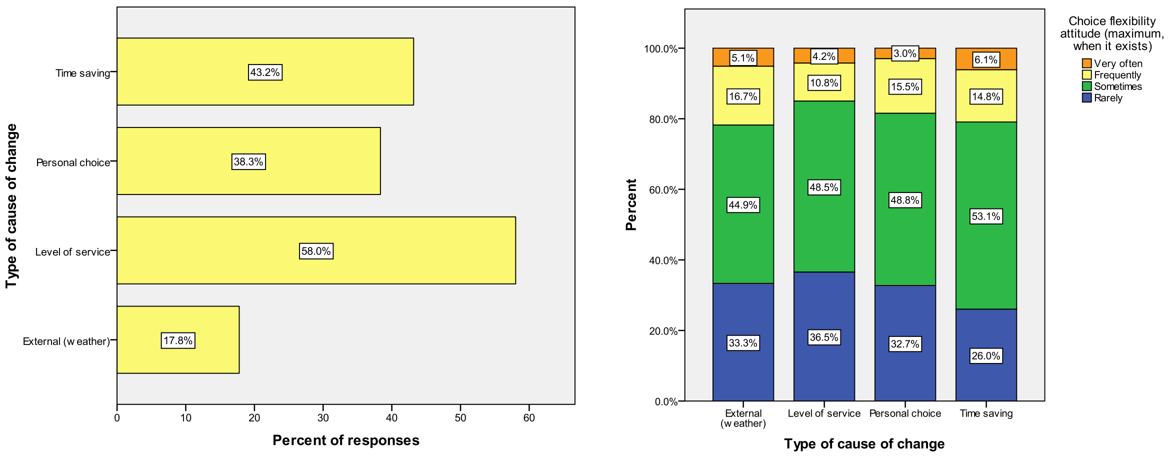


Figure 10 – Cause of change – (a) Occurrence, (b) Frequency of changes

Insights into choice flexibility

Existence of the attitude to change

In the following the existence of an attitude to change has been defined for those respondents who have alternative options in at a least one of the four dimensions considered for the flexibility of choices (i.e., respondents whose answer in at least one of the questions presented in Figure 5-Figure 8 is not “Never because you don’t have alternatives”). It is represented by a binary variable, *Existence of attitude*, which assumes the value “No” if in each choice dimension the answer is “Never, even though you could”, and “Yes” otherwise. This results in 358 respondents who do (at least to some degree) have a positive attitude to travel on hyperpaths, whereas 43 respondents will only consider a single path to their destination.

The existence of an attitude to change has been studied in relation to four sets of trip attributes:

- The knowledge the traveller has got concerning service times
- The level of the transit service used, in terms of usual crowding of vehicles (*Usual vehicle congestion*) and reliability of travel time (*Arrival time predictability*)
- Some intrinsic characteristics of trips: The duration of the trip on average (*Average trip time*); the potential reduction of the trip duration, calculated according to the below expression

  $$100 \cdot \frac{\text{avTT-minTT}}{\text{avTT}}$$

  where minTT and avTT are respectively the minimum and the average duration of the trip declared by each respondent (*Potential trip time reduction*); the potential excess trip time (*Potential trip time increase*), equal to

  $$100 \cdot \frac{\text{maxTT-avTT}}{\text{avTT}}$$

where maxTT is the maximum time a trip can take; the minimum number of changes involved in a given trip (*Min number of changes*).

Note that potential trip time variations and arrival time predictability are semantically different concepts though related to each other: A given variation can be more or less easily anticipated, whereas in an unpredictable network the variations of trip duration can be low or high. If variability is low, predictability becomes less important and so it can be expected to be overrated by respondents. Actually Figure 11 shows that high values of predictability (corresponding to low values of the variable *Arrival time predictability*) and low variations tend to be associated but the association is weak. Of course physical restraints make TT potential reductions lower than increases. In the model possible TT reductions are introduced as the regret usually experienced by travellers because their trips on average last more than the minimum trip time, whereas possible TT increases represent fear of further excess trip time.

Correlations

			Arrival time predictability	Possible TT reduction	Possible TT increase
Kendall's tau_b	Arrival time predictability	Correlation Coefficient	1.000	.127**	.218**
		Sig. (2-tailed)	.	.001	.000
		N	643	441	602
	Possible TT reduction	Correlation Coefficient	.127**	1.000	.314**
		Sig. (2-tailed)	.001	.	.000
		N	441	456	456
	Possible TT increase	Correlation Coefficient	.218**	.314**	1.000
		Sig. (2-tailed)	.000	.000	.
		N	602	456	641

**. Correlation is significant at the 0.01 level (2-tailed).

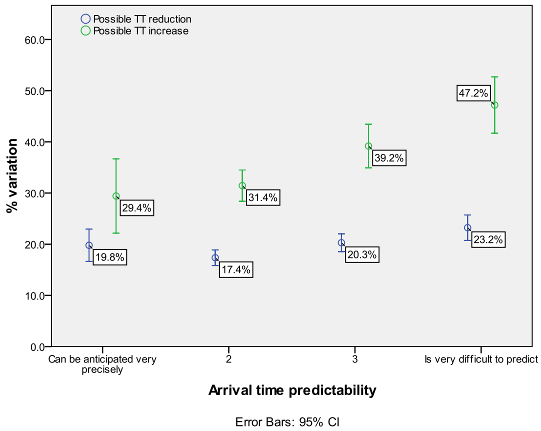


Figure 11 – Arrival time predictability and variations of trip times

- Two elements of the route choice process linked to the value of the trip for the respondent: The purpose of the trip (*Trip purpose*) and the importance of arriving on time (*Importance of punctuality*).

Regression models do not prove causality but correlation and the choice of the dependent variable is sometimes arbitrary. In the models of the attitude to change, this is the case at least for the variable expressing the intrinsic characteristics of the trips: e.g., we assume the implication "trip time variation ⇒ attitude to change" but also "attitude to change ⇒ trip time variation" is of course feasible, and consistent with the hyperpath argument that by defining complex strategies one gains the potential to reduce the trip time. The choice here is motivated by the interest in the flexibility of user decisions.

The existence of the attitude to change has been interpreted by binary logistics models (Hosmer et al., 2000 can be referred to for statistics used in this section, when not differently specified) whose characteristics are summarized in Table IV. Every category of the nominal variable *Trip purpose* has been compared to the "Commuting" one. The parameters of the categorical variables have been estimated with the reverse Helmert contrast method

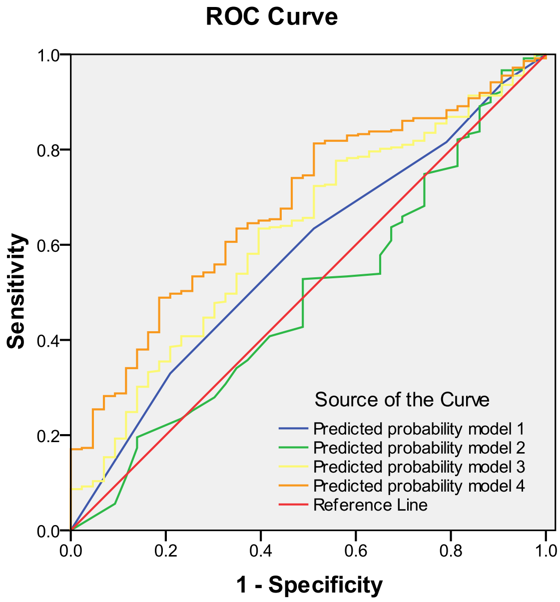
(Stevens, 2002), in which every category except the first is compared to the average effect of the previous ones.

The omnibus tests show that all the models are significantly different from the one including only the intercept, but the Hosmer and Lemeshow test – commonly recommended for testing the overall fit of binary logistic models when continuous covariates are included as independent variables – finds that only the model considering all the variables fit adequately data. The odds in favour of the existence of the attitude to change is 8.32, so it can be expected to be almost impossible identifying models which can produce a classification of cases definitely better than that obtainable by chance. In fact when a threshold of 0.500 is assumed, models 3 and 4 have a classification rate approximately equal to that produced by chance – calculated by the proportional reduction in error method or by the proportional by chance one – whereas models 1 and 2 are less able to replicate respondents' behaviour. Higher total classification rates are reached by reducing the capacity of the models to assign correctly cases of not existence of the attitude. The ROC curves (Figure 12) show that models 1 and 2 do not perform significantly better than "flipping a coin", whereas models 3 and 4 have a better, even though not high, capacity of classifying cases correctly; model 4 is slightly better than model 3, but their difference cannot be proven at a 95% level of confidence. The Cox and the Snell and Nagelkerke's approximations to R^2 corroborate the results on the quality of models, by indicating that the association between the model and the actual data becomes stronger when new variables are introduced and that models 3 and 4 are pretty equivalent. The results on the goodness of fit of the logistic models can be interpreted as pointing to the fact that the existence of the attitude to change cannot be adequately explained only by the level of user knowledge and by the service condition, but it is related to the characteristics of the trip, intrinsic and purpose-related. In other words, it cannot be assumed that a traveller will introduce changes in his itinerary just because he knows (or does not know) the available services, or because the level of operations is not satisfactory, but the choice is related also to the duration of the trip, the number of compulsory changes, the purpose of travelling and the value of arriving on time [result 8].

From model 4, the most complete and informative, it can be seen that no significant difference in having an attitude towards change is found for different knowledge about departure time and level of services, and for different trip purposes. The first two groups of variables have significant Exp(B)'s in models 1 and 2 but they become statistically not significant in models 3 and 4, from which it can be inferred that the variables added in the more complex models shadow their role in determining travellers' behaviour. The effect of the average trip duration and of the fear of larger excess trip time (*Potential trip time increase*) cannot be proven at 95%. A weak effect has the regret experienced on average (a 1% increase in *Potential trip time reduction* increases the odds of having an attitude to change by 5.5%) [result 9]. The minimum number of changes has a positive effect (1 change more is equivalent to an increase of odds by 63.9%), which means a positive feedback of the necessity of changing onto the attitude to do the same even when not compulsory [result 10]. *Importance of punctuality* has two significant levels out of four showing that the higher the value of arriving on time, the more travellers tend to introduce optional changes in their trips. Further investigation is needed to confirm the last result. In fact the same models in Table IV have been estimated using simple contrasts, which test the difference between the reference attribute and each other attribute; for every ordinal variable, the lowest category (i.e., that in the first row of each variable in the table) has been chosen as the reference category. Results are obviously comparable to those reported, but show that sometimes the values of the parameters of single attributes within variables are not consistent. In particular in model 4 the parameters of *Importance of punctuality* increase from the attribute "2" to the attribute "4" but decrease from "4" to "5" [result 11].

Table IV – Existence of attitude to change

Existence of choice flexibility (binary logistic model)										
Dependent: Attitude towards change: no = 0 (43 cases), yes = 1 (358)										
			Model 1		Model 2		Model 3		Model 4	
			Exp(B)	Sig.	Exp(B)	Sig.	Exp(B)	Sig.	Exp(B)	Sig.
Knowledge	Knowledge about line departure times [ord*]	Only the time you have to leave		**0.000**		**0.000**		0.479		0.419
		The line frequency at your starting stop/station	1.253	0.442	1.278	0.462	1.570	0.315	1.496	0.385
		The line frequencies at each transfer point of your trip	0.318	**0.000**	0.311	**0.001**	0.743	0.552	0.813	0.700
		The complete timetable only at your starting stop/station	0.683	0.155	0.785	0.409	0.657	0.278	0.633	0.260
		The complete timetable at each transfer point along your trip	0.044	**0.000**	0.112	**0.000**	0.576	0.304	0.481	0.181
Service condition	Usual vehicle congestion [ord]	You can always find a seat				**0.000**		0.897		0.825
		Sometimes you have to stand			3.209	**0.001**	0.974	0.954	1.036	0.939
		You always have to stand			0.955	0.883	1.072	0.874	1.155	0.755
		Sometimes you can't even get onto the first vehicle			0.616	0.115	0.717	0.447	0.676	0.376
	Arrival time predictability [ord]	1 = Can be anticipated very precisely				**0.000**		0.529		0.427
		2			5.618	**0.000**	0.680	0.532	0.676	0.543
		3			1.744	0.066	0.586	0.217	0.548	0.181
		4 = Is very difficult to predict			1.091	0.775	0.619	0.272	0.570	0.209
Trip characteristics - Intrinsic	Average trip time [num]						0.994	0.422	0.993	0.343
	Potential trip time reduction [num]						1.050	**0.002**	1.055	**0.001**
	Potential trip time increase [num]						1.017	**0.037**	1.016	**0.075**
	Min number of changes [num]						1.647	**0.010**	1.639	**0.015**
Trip characteristics - Meaning for the traveller	Trip purpose [nom]	Commuting								0.713
		Work-related							2.121	0.320
		Personal/family businesses							1.174	0.807
		Other activities							1.449	0.523
	Importance of punctuality [ord]	1 = Not important								**0.061**
		2							1.744	0.440
		3							4.073	**0.045**
		4							2.597	**0.050**
		5 = Important							1.194	0.632
Omnibus Tests of Model Coefficients		Chi-square	90.453		135.868		281.855		291.456	
		df	4.000		10.000		14.000		21.000	
		Sig.	**0.000**		**0.000**		**0.000**		**0.000**	
Hosmer and Lemeshow Test		Chi-square	309.480		147.729		16.635		9.934	
		df	3.000		8.000		8.000		8.000	
		Sig.	0.000		0.000		0.034		**0.270**	
Correct classification (cut value = 0.500)						Δ% pre. mod.		Δ% pre. mod.		Δ% pre. mod.
		No (%)	20.9		14.0	-33.3%	2.3	-83.3%	2.3	0.0%
		Yes (%)	81.6		83.2	2.1%	100.0	20.1%	99.2	-0.8%
		Total (%)	75.1		75.8	1.0%	89.5	18.1%	88.8	-0.8%
Model Summary						Δ% pre. mod.		Δ% pre. mod.		Δ% pre. mod.
		Cox & Snell R Square	0.202		0.287	42.3%	0.505	75.7%	0.517	2.3%
		Nagelkerke R Square	0.269		0.383		0.673		0.689	



Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval: Lower Bound	Asymptotic 95% Confidence Interval: Upper Bound
Predicted probability model 1	.575	.045	.110	.487	.662
Predicted probability model 2	.487	.047	.777	.395	.578
Predicted probability model 3	.621	.044	.010	.535	.707
Predicted probability model 4	.682	.039	.000	.605	.758

The test result variable(s): Predicted probability model 1, Predicted probability model 2, Predicted probability model 3 has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Figure 12 – Fit of the different models of the existence of the attitude to change

Travellers who do not draw on any kind of external information – reasonably because they have a knowledge of the system which is sufficient to move over it without any help – are those with the highest rate of changes (Figure 12) [result 12]. Information gathered during the journey is associated with more changes than that collected only at home.

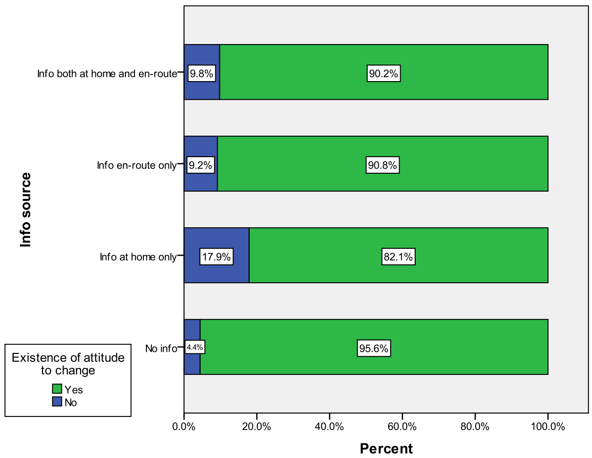


Figure 13 - Existence of the attitude to change and source of information

Level of attitude to change

The inclination towards changing has been measured, for people who have *Existence of attitude* = “Yes”, through the categorical variable *Level of attitude to change*, with 4 attributes ranging from “Rarely” to “Very often”. The value is equal to the maximum frequency of changes declared in questions shown from Figure 5 to Figure 8. The relation of *Level of attitude to change* with the same groups of explanatory variables used to model the existence of the attitude has been studied through the categorical regression procedure

(CATREG) implemented in SPSS (SPSS, 2009). CATREG makes use of optimal scaling techniques to find the best-fitting model. The technique is appropriate given that the main aim of the analysis is not forecasting but identifying significant interactions, and it has been preferred to a traditional multinomial logit regression because of the shortage of data for a logit estimation of the more complete models.

Categorical variables, including the dependent, have been scaled as second-degree monotonic splines with two interior knots; for the ordinal ones the optimal scaling preserves the order of categories. The numeric variables have been discretized in 7 equally distanced categories.

The ANOVA shows that all the models except the one including only *Knowledge* are significantly different from the one in which all the (transformed) variables have null regression coefficients (Table V). For the same models, R^2 is low but acceptable given the nature of the variables and the aim of the analysis. However adjusted R^2 shows that the slight improvement in explaining the (transformed) observed variance of the dependent which is produced by adding *Trip purpose* and *Importance of punctuality* can be due to an overfitting of the model, so model 3 can be considered better than model 4.

Assuming a 95% level of confidence, *Usual vehicle congestion* is significant both not considering (model 2) and considering (model 3) trip-specific characteristics. Moreover Pratt's measures of importance (Thomas et al., 1998) in both models agree in assigning the variable the first place among the explanatory factors. Therefore the influence of vehicle crowding on the existence of the attitude to change is not proven, but if a passenger has such an attitude, the frequency of changes depends on it [result 13]. The standardized β's and the quantifications (not reported for sake of brevity) point to a positive effect of the congestion of vehicles on the frequency, showing that the three lowest level of vehicle crowding are not really distinguishable, whereas a major change in travellers' behaviour happens when carriers are usually so packed that sometimes it is impossible to board the first vehicle arriving.

Table V – Level of attitude to change

Level of choice flexibility (categorical regression model)																	
Dependent: Level of attitude to change																	
		Model 1				Model 2				Model 3				Model 4			
		rarely (141cases), sometimes (168), frequently (47), very often (16)				rarely (134 cases), sometimes (164), frequently (45), very often (16)				rarely (79 cases), sometimes (112), frequently (30), very often (12)				rarely (79 cases), sometimes (111), frequently (30), very often (12)			
		Std β	Sig.	Pratt's	Tole. af.	Std β	Sig.	Pratt's	Tole. af.	Std β	Sig.	Pratt's	Tole. af.	Std β	Sig.	Pratt's	Tole. af.
Knowledge	Knowledge about line departure times [ord*]	0.103	0.268	1.000	1.000	0.069	0.454	0.085	0.998	0.125	0.258	0.157	0.974	0.112	0.414	0.120	0.965
Service condition	Usual vehicle congestion [ord]					0.225	**0.000**	0.886	0.991	0.168	**0.045**	0.348	0.887	0.144	0.161	0.242	0.875
	Arrival time predictability [ord]					0.033	0.719	0.029	0.990	-0.064	0.627	-0.007	0.875	-0.048	0.700	-0.005	0.861
Trip characteristics - Intrinsic	Average trip time [num]									-0.119	0.213	0.158	0.824	-0.137	0.177	0.158	0.818
	Potential trip time reduction [num]									0.137	**0.085**	0.270	0.826	0.137	**0.081**	0.226	0.826
	Potential trip time [num]									0.012	0.892	0.014	0.771	0.009	0.921	0.008	0.766
	Min number of changes [num]									0.068	0.389	0.060	0.908	0.070	0.430	0.054	0.905
Trip characteristics - Meaning for the traveller	Trip purpose [nom]													0.125	**0.081**	0.123	0.909
	Importance of punctuality [ord]													0.093	0.583	0.074	0.935
ANOVA	Sig.	0.142				**0.002**				**0.009**				**0.043**			
Model summary								Δ% pre. mod.				Δ% pre. mod.				Δ% pre. mod.	
	Multiple R	0.103				0.241		135.1%		0.303		25.8%		0.330		8.6%	
	R Square	0.011				0.058		452.8%		0.092		58.2%		0.109		18.0%	
	Adjusted R Square	0.005				0.042		716.2%		0.055		31.5%		0.047		-15.7%	
* ord = ordinal variable, num = numeric variable, nom = nominal variable																	

From Figure 14 it appears that the frequency of changes increases with the richness of the set of sources of information, i.e. passing from "No info" to "Info both at home and en-route".

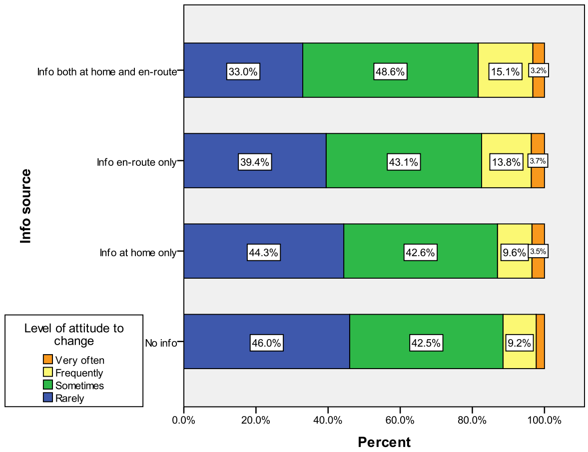


Figure 14 - Level of the attitude to change and source of information

DISCUSSION OF RESULTS

Before discussing the results of this explorative study in detail it is important to remind that our sample is clearly biased towards the young, male expert public transport users in a few selected countries: Though this limits the general validity of our results, we believe that some important observations can be made on route choice flexibility.

We observe that trip itineraries are not fixed in most cases supporting the argument to model route choice as a "hyperpath" [see result 4]. However the actual behaviour of transit users presents some differences from what is supposed in most transit hyperpath route choice models: Firstly, the tendency to consider more lines at a given stop is not so pronounced; this is equivalent to say that the attractive choice sets at stops tend to be made up of just a single line. It might be rather the stop or platform choice of passengers that should be modelled as a hyperpath. Secondly, but connected to the first point, in particular the most frequent kind of change concerns the departure stop/station, whose choice is often ignored by models [for both the remarks see result 6 and result 7]. This can be interpreted as an indication that usually transit network representations are assumed which are not consistent with travellers' mental maps. A greater consideration of the importance of anchor points in transit modelling seems to be endorsed by the results of the survey.

Very few respondents have explicit knowledge about service timetables and frequencies at all the transfer points of their reported trips [see result 2], even though these are usual trips and only rarely entail more than two changes [see result 1]. However, this does not prevent people from modifying their itineraries quite often [see result 5]. Depending on the complexity of the network and on the effects of learning by repetition (reinforcement), such modifications might be different from those derived from the assumption of perfect information: In simple networks it is likely that the actual choice sets are not very dissimilar from the perfect information ones because of the existence of few alternatives, whereas this difference could become relevant in networks with frequent and overlapping services. A possibly counterintuitive hint on the role of reinforcement in the way transit users deal with network representation comes from the relation between the existence of the attitude to change and the information. One would expect that more information provides the traveller with a larger set of options. Instead in the models in which *Knowledge* has significant categories, the travellers with more information on service departure times show a weaker attitude to change [see Table IV]. The results further suggest that those who (perceive to) have an adequate knowledge of the network and do not use any information sources are those most likely to change their route [see result 12]. Taking these results together, one might conclude that information and day-to-day learning tends to lead to a rather fixed, simpler route set considered by travellers. In other words, information and reinforcement could lead to a reduction of the complexity of the actual choice set, i.e. the sub-set of the optimal lines which is taken into consideration by the traveller for a specific trip, rather than to its enlargement. The argument could be reversed by saying that less information is needed when systems are simpler, for instance because fewer alternatives exist or because services are very punctual so travellers do not need to consider alternatives. Note that the interpretation that people without information change more because their behaviour is someway random seems difficult to be sustained because replies concern habitual trips. Further research on the demand for and the use of information in routine transit trip is needed to better understand and model the role of information on passenger mental representation of networks.

Of course a lack of explicit and/or implicitly accumulated knowledge can be compensated by relying on information sources, but the information systems in use at the moment do not foster rational adaptive behaviour, because they often only consider only a partial and deterministic network [see result 3]. Most information systems do not assist travellers in the calculation of shortest paths/hyperpaths (timetables and displays), or they provide travellers with suggestions on alternative single paths – which assume no variance in service times or frequencies – and cannot be updated according to the real time conditions (on line journey planners at home). The actual behaviour can be expected to become more similar to the rational one assumed in transit models as a consequence of an increasing diffusion of mobile phones in the field of travel information: In fact at the moment they can provide the access to on-line journey planners while travelling and, in the near future, they will supply navigation assistance on their own. The positive effect on rational choice can be strengthened by the underground access to mobile signal, which, when not yet in place, is increasingly in the plans of ICT companies and transport operators. An in-depth analysis of the effects of not complete information and of heuristics used to deal with it is required also to evaluate the potential benefits of the diffusion of transit oriented navigation assistance devices.

The models built to explain the existence of an attitude to change show that it correlates in a significant way to the intrinsic characteristics of the trip (duration on average, expected and feared excess travel, minimum number of changes) and to the meaning of the trip itself to the traveller (purpose and importance of punctuality) [see result 8]. A positive effect on the existence of the attitude is proved for the expected excess trip time [see result 9], the minimum number of changes [see result 10] and, with some caveats, the relevance of on-time arrivals [see result 11]. Such findings contradict the assumption usually underpinning transit modelling that the travel behaviour is irrespective of the trip characteristics (e.g. in

determining a hyperpath a line is added to a choice set even if this causes a very small reduction of expected travel time in an already short trip) and supports the development of models considering expectations and regret, fuzzy decision criteria, and multiclass users.
The minimum number of changes is an indicator of the complexity of a trip and it is reasonable to assume that its positive influence onto the attitude to change is due mainly to the fact that more compulsory changes mean more chances of not compulsory changes. But given that the investigated dimensions of changes include also type of changes not related to intermediate stops (i.e. changing departure point and changing an already boarded line), the finding can admit also another explanation: The existence of “dynamic” travellers, who become "fitter to changes" because of "training". It is a suggestive hypothesis worth being tested, which does not contrast with the widely accepted idea that changes are associated with costs, because it has to do with an attitude which can be more or less exerted depending on the characteristics of the system used by a traveller.
The link between vehicle overcrowding and higher frequency of change [see result 13] is expected and calls for the introduction of seat availability information in route choice and assignment models. As with other tentative conclusions in this discussion one might however qualify this argument by the observation that the most crowded cities in our sample are also the ones with the highest number of route options. This and further issues will be addressed in future data analysis considering the geographical distribution of replies.

CONCLUSIONS

Route choice models considering hyperpaths are often applied to represent the transit user decision making process, especially in systems with high frequency and overlapping lines. The idea that travellers choose among a set of attractive lines looks realistic, yet the existing research on route choice seems to suggest that the behavioural model on which the definition of the attractive set and the choice of the service rely could be too simplistic.
The international survey described in this paper has been carried out to start dealing with the issue, with the aim of identifying questions worth being addressed in further behavioural research to improve transit route choice and assignment models. The analysis presented in the paper highlights some relevant problems and in brackets some very tentative answers based on our analysis:

- Are cognition maps of transit networks, and therefore route choice, based on anchor points more than on lines? (Our initial analysis suggests that passengers choose their starting station carefully)
- Does learning through repetition lead to an enlargement or to a shrinkage of the traveller knowledge base? (We suggest rather shrinkage)
- What heuristics are used by travellers to move under imperfect information? What impact could the provision of real time information and navigation assistance have on transit user behaviour?
- How important are expectations and regret in evaluating alternatives? Do travellers really decide on the basis of differences of expected travel times, however small they may be? (We suggest that regret is indeed an important aspect in transit route choice)

- Should other components be added to trip time in the cost function of transit users, e.g. familiarity with changing and seat availability? (Our analysis suggests yes to both examples).

Not all the information collected by the survey has been explored yet. Future work will be targeted to specify and to find better support answers to some of the previous questions. In particular new insights are expected by considering the origin of respondents. Such information can help understanding the role of the features of public transport systems on the user behaviour. The section investigating preferences in hypothetical route choice scenarios will shed light on the attitude of travellers to deal with expected travel times in presence of waiting times, and to test whether the decision making process supposed by Spiess and Florian is realistic.

REFERENCES

Conover, W. J. (1999). Practical nonparametric statistics. Wiley, New York.

De Cea, J. and E. Fernandez (1993). Transit Assignment for Congested Public Transport Systems: An Equilibrium Model. Transportation science, 27, 133.

Fonzone, A. and M. G. H. Bell (2010). Bounded Rationality in Hyperpath Assignment: the Locally Rational Traveller Model. Proceedings of TRB 89th Annual Meeting, Washington, D.C.

Gentile, G., S. Nguyen and S. Pallottino (2005). Route choice on transit networks with online information at stops. Transportation science, 39, 289-297.

Golledge, R. (1997). Defining the criteria used in path selection. In: Activity-based approaches to travel analysis (D. Ettema and H. J. P. Timmermans, ed.), pp. 151-169. Pergamon, [S.l.].

Golledge, R. G. and R. J. Stimson (1997). Spatial behavior : a geographic perspective. Guilford Press, New York.

Golledge, R. G. and T. Garling (2001). Spatial Behavior in Transportation Modelling and Planning. In: Transportation and Engineering Handbook (K. Goulias, ed.).

Hall, R. W. (1986). The Fastest Path through a Network with Random Time-Dependent Travel Times. Transportation science, 20, 182.

Hamdouch, Y. and S. Lawphongpanich (2008). Schedule-based transit assignment model with travel strategies and capacity constraints. Transportation research. Part B, Methodological, 42, 663-684.

Hosmer, D. W., S. Lemeshow and E. D. Cook (2000). Applied logistic regression. Wiley, New York.

Lyons, G. (2006). The role of information in decision-making with regard to travel. IEE proceedings, 153, 199-212.

Nguyen, S. and S. Pallottino (1988). Equilibrium traffic assignment for large scale transit networks. European journal of operational research, 37, 176-186.

Nokel, K. and S. Wekeck (2009). Boarding and alighting in frequency-based transit assignment. 88th TRB Annual Meeting, Washington, D.C.

Papinski, D., D. M. Scott and S. T. Doherty (2009). Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS. Transportation research. Part F, Traffic psychology and behaviour, 12, 347.

Schmöcker, J. D., A. Fonzone and M. G. H. Bell (2009). Assignment under consideration of seat availability: An application to London's underground network. Mobil.TUM 2009 - International Scientific Conference on Mobility and Transport, Munich.

Seaborn, C. (2009). Analyzing Multimodal Public Transport Journeys in London with Smart Card Fare Payment Data. Transportation research record, 2121, 55-62.

Shimamoto, H., F. Kurauchi, J. D. Schmöcker and M. G. H. Bell (2010). Transit Assignment Model Considering the Correlation of Vehicles' Arrival. Proceedings of TRB 89th Annual Meeting, Washington, D.C.

Spiess, H. and M. Florian (1989). Optimal strategies. A new assignment model for transit networks. Transportation research. Part B, Methodological, 23, 83-102.

SPSS (2009). SPSS Statistics 17.0 - Algorithms. SPSS Inc., Chicago, Illinois.

Stevens, J. (2002). Applied multivariate statistics for the social sciences. Lawrence Erlbaum Associates, Mahwah, N.J.

Thomas, D. R., E. Hughes and B. D. Zumbo (1998). On Variable Importance in Linear Regression. Social Indicators Research, 45, 253-275.

Wolpert, R. (1964). The decision process in a spatial context. Annals of the Association of American Geographers, 54, 537-558.